ORIGINAL ARTICLE



Fully automatic detection of renal cysts in abdominal CT scans

Neta Blau¹ · Eyal Klang² · Nahum Kiryati¹ · Marianne Amitai² · Orith Portnoy² · Arnaldo Mayer²

Received: 18 November 2017 / Accepted: 28 February 2018 / Published online: 15 March 2018 © CARS 2018

Abstract

Purpose Simple renal cysts are a common benign finding in abdominal CT scans. However, since they may evolve in time, simple cysts need to be reported. With an ever-growing number of slices per CT scan, cysts are easily overlooked by the overloaded radiologist. In this paper, we address the detection of simple renal cysts as an incidental finding in a real clinical setting.

Methods We propose a fully automatic framework for renal cyst detection, supported by a robust segmentation of the kidneys performed by a fully convolutional neural network. A combined 3D distance map of the kidneys and surrounding fluids provides initial candidates for cysts. Eventually, a second convolutional neural network classifies the candidates as cysts or non-cyst objects.

Results Performance was evaluated on 52 abdominal CT scans selected at random in a real radiological workflow and containing over 70 cysts annotated by an experienced radiologist. Setting the minimal cyst diameter to 10 mm, the algorithm detected 59/70 cysts (true-positive rate = 84.3%) while producing an average of 1.6 false-positive per case.

Conclusions The obtained results suggest the proposed framework is a promising approach for the automatic detection of renal cysts as incidental findings of abdominal CT scans.

Keywords Kidney segmentation · Cyst detection · CT · FCN · CNN

Introduction

Simple renal cysts are benign and asymptomatic fluid collections in the kidney, commonly occurring in aging subjects. They are characterized by an approximately spherical shape and homogeneous fluid content, without internal septations. They are usually found incidentally in patients undergoing abdominal imaging for other reasons [1], especially in computed tomography (CT). In most cases, no treatment is required. However, cysts must be reported by the radiologist and sometimes require follow-up as they may evolve in time. With the progress of CT technology, the number of slices per scan is constantly growing, and cysts are often overlooked by the overloaded radiologist. In this paper, a fully automatic

method is presented for the detection of simple renal cysts as an incidental finding in a real clinical setting.

Fully automatic segmentation of the kidneys given 3D CT scans is a prerequisite to our objective. Automatic detection and segmentation of organs remain challenging due to the high inter-patient variability of the organs' shapes and locations. Recently, several studies utilized fully convolutional neural networks (FCN) for end-to-end segmentation in medical images, and specifically in CT scans. Zhou et al. [2] trained an FCN to segment 19 different organs in 2D slices from all three anatomical planes of a CT scan. The resulting segmentations are later fused by majority voting into a 3D segmentation. State-of-the-art performance was reported especially for the segmentation of larger organs such as the liver and the heart. Hu et al. [3] proposed a method for liver, spleen and kidney segmentation that employs a 3D FCN followed by a time-implicit multi-phase level-set algorithm for segmentation refinement. Reported accuracy compares to state-of-the-art methods with shorter running time. Christ et al. [4] applied cascaded fully convolutional networks to segment the liver and possible lesions. The networks were combined with a dense 3D conditional random field and provided satisfactory results. A similar approach



Department of Electrical Engineering, Tel-Aviv University, Ramat-Aviv, Israel

Diagnostic Imaging Institute, Sheba Medical Center, Affiliated with Sackler School of Medicine, Tel Aviv University, Tel-Aviv, Israel

was implemented by Dou et al. [5] with a 3D FCN. While the above algorithms segment healthy organs well, our task is harder considering the strong deformations that kidneys may sustain in the presence of large cysts.

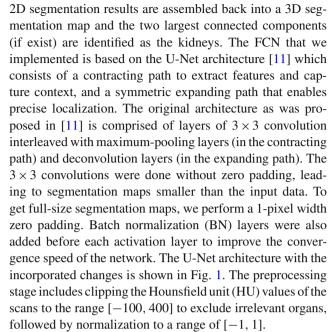
Beyond the segmentation of the kidneys, the problem of detecting simple renal cysts remains to be addressed. Due to their random origins, cysts vary significantly in size, location and quantity. Owing to the homogenous fluid consistency of the cysts, one may think that simple thresholding of the kidney area can suffice for cyst detection. However, being part of the urinary system, other fluid collections are normally present in the kidneys, making it difficult to distinguish renal cysts from healthy functioning parts of the kidneys such as the renal pelvis and renal calyces. Few studies were published on automatic renal cyst and/or lesion detection. Summers et al. [6] employed thresholding and morphological operations to segment renal cysts and solid tumors and estimate their volumes. However, the work focused on the quantification of the lesions and detection rates were not reported. Battiato et al. [7] proposed a method to refine a cyst segmentation given an initial boundary marked by the radiologist. The refinement is done in the annotated slice using a bank of filters and is then propagated to the other slices. However, in this method, the cysts must be detected by the users, and their intervention is required for the segmentation. Piao et al. [8] proposed to segment cysts using fuzzy C-mean clustering but did not report any quantitative results. An algorithm proposed by Badura et al. [9] employs thresholding and morphological operations for the detection of cyst candidates. The candidates are then classified into cysts and non-cysts based on 3D shaperelated features. This approach produced promising results, but the algorithm was evaluated on a fairly small dataset of 16 patients with 25 cysts. Another solution, proposed by Liu et al. [10], relies on geometric changes on the kidney surface for cyst candidate detection. Again, shape features are utilized for the classification of the candidates. This algorithm was indeed validated on a large dataset but is limited to detecting only exophytic lesions, and the reported false-positive rate (15 per patient) was high.

We present a fully automatic algorithm for renal cyst detection based on a neural framework. The algorithm consists of three parts: kidney segmentation, selection of cyst candidates and binary classification of the candidates. The output of the system is the 3D center position for each detected cyst. Performance is validated both qualitatively and quantitatively on a large dataset, with promising results.

Methods

Kidney segmentation

The 2D axial slices of the CT scan are fed into an FCN which classifies each voxel as *kidney* or *non-kidney*. The



To improve the segmentation robustness to abnormal kidneys, any renal cyst or other pathology observed in the training set was annotated as *kidney* as well.

The cost function we use for training the FCN is the softmax log-loss given by:

$$cost = -\sum_{i,j} log \frac{exp\left(x_{ij}^{c}\right)}{\sum_{d=1}^{D} exp\left(x_{ij}^{d}\right)}$$
(1)

where $c \in [1...D]$ is the ground truth class for pixel (i, j) and x is the predicted score matrix. The term inside the log function converts the predicted score into a value between [0,1]. This value represents the probability for pixel (i, j) to belong to the correct class. A high probability value, close to 1, will result in a low cost while a low probability value will result in a high cost. In our case there are only two classes, kidney and non-kidney, and the described cost function represents our objective to determine the correct label for each pixel with high probability.

Identification of cyst candidates

A cyst candidate is the center of a fluid collection, physically attached to one of the kidneys. To locate cyst candidates, we generate two 3D Euclidian distance maps by applying a 3D distance transform (DT) [12]. Since the spacing between the slices of a CT scan may be different than its pixel spacing, an anisotropic DT is necessary.

In the first map, the value of each voxel is its distance from the closest *non-fluid* voxel, where a fluid voxel is one with HU in the range [-50, 80]. As demonstrated in Fig. 2 (top), *non-fluid* voxels obtain zero value. A Cyst seed-point,



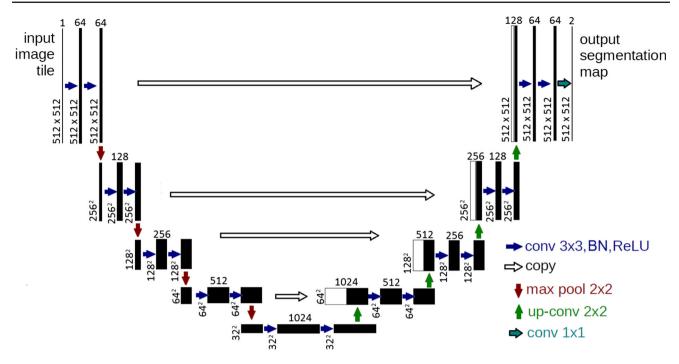


Fig. 1 The kidney segmentation FCN based on the U-Net architecture. The figure is adapted from [11]

or, in short seed, is defined as a point located at a regional maximum of the distance from non-fluid map, denoted by DNF in the sequel. These are, by definition, the voxels at the center of a fluid collection and appear hyperintense in DNF. For simple renal cysts, which are approximately spherical, the seed's distance to non-fluid would be the sphere's radius. We denote the seed's distance to non-fluid as $D_{\rm nf}$, and refer to it as the estimated radius of the seed. To reduce the number of potential seeds and considering the scan's resolution, seeds with estimated radius smaller than the spacing between slices are eliminated. Still, as can be seen from the DNF in Fig. 2 (top), many neighboring seed candidates (red crosses) are detected which are not the approximate center of a cyst. For this purpose, a second distance map is defined to increase seeds specificity.

In the second distance map, DK, the value of each voxel is its distance from the closest kidney voxel, denoted $D_{\rm k}$. We define the adjacency of a seed to the kidney as $D_{\rm nf}-D_{\rm k}$. For spherical objects, a nonnegative adjacency indicates the object and the kidney share a common wall. A negative value indicates they are separate. If a perfect segmentation of the kidneys was available, cyst candidates would be seeds with a nonnegative adjacency value. However, to account for cases of under-segmentation of the kidneys, we decided to loosen the adjacency requirement. Missing a voxel at the edge of the kidney leads to an error of 2.5 mm in the worst case, since it corresponds to the largest dimension of the voxels, which is in z direction. Therefore, the adjacency threshold was set to $\theta_{\rm adj}=-2.5$ mm.

The process of seed selection is further illustrated in the enlarged view of Fig. 2 (bottom), for 3 seed candidates (yellow ROI) extracted from the DNF. The same enlarged ROI is shown for the DK. The values correspond to the distance from *non-fluid* (left map) and distance to kidney (center map) for the 3 seed candidates. The subtraction between the maps gives the adjacency (right map) for the 3 candidates. Only one candidate seed is retained (green point, adjacency = 14) as it surpasses the threshold value $\theta_{adj} = -2.5$ mm.

Classification of cyst candidates

Following the detection of cyst candidates, we wish to determine which of them are indeed cysts. The main steps for the proposed cysts candidate classification algorithm are shown in Figs. 3. and 4.

Data preparation

We note that the cysts are radially symmetric and hence their 2D projection has similar properties in any orientation. Capitalizing on this notion, patches of the 2D projections of each candidate on three orthogonal planes (axial, sagittal and coronal) are used to construct a 3-channel image. This representation of a volume-of-interest is known as a 2.5D image [13]. The patch size is 32×32 pixels and it is resampled from the original CT scan according to the seed's estimated radius so that the suspected object will occupy the central 10×10 pixels of the patch. Resampling is done using bilinear inter-



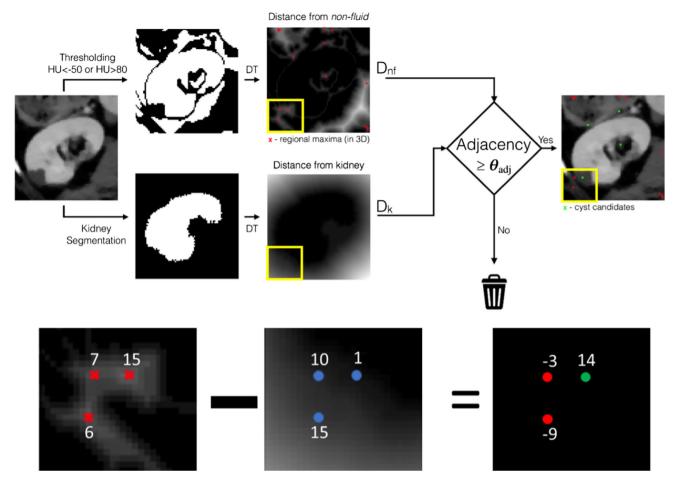


Fig. 2 (Top) Main steps of the cyst candidates' selection algorithm; (bottom) Enlarged view of the seed selection process for three candidates. (Yellow ROI) Extracted from the DNF. The same enlarged ROI is shown for the DK. The values correspond to the distance from non-fluid

(left map) and distance to kidney (center map) for the 3 candidates. The subtraction between the maps gives the adjacency (right map) for the 3 candidates. Only the green seed is retained since its adjacency (= 14) is $> \theta_{\rm adj}$

polation, after which the patches are smoothed by a Gaussian filter ($\sigma = 0.5$, kernel size = 3×3) and clipped to HU range [-50, 80].

To enhance the number of cyst positive examples, data augmentation is also performed by rotating and flipping the original labeled examples, leading to a positive versus negative examples ratio of about 1:8.

CNN classification

The 2.5D images of the candidates are fed into a small convolutional neural network, composed of three layers of 3×3 convolutions and 2×2 max pooling followed by two fully connected (FC) layers which outputs the probability of each seed being a cyst (see Fig. 5). Seeds with probability greater than 0.5 are classified as cysts.

As a final postprocessing step, cysts that are contained in another cyst, i.e., their seed is within the estimated radius of another seed, are considered redundant and eliminated.



Kidney segmentation

The dataset for kidney segmentation included 46 abdominal CT scans with contrast agent, taken at the portal phase of enhancement. The CT cases were randomly selected within a one-month period and the kidneys were manually annotated by senior radiologist EK. As mentioned in the "Kidney segmentation" section, any renal cyst or other pathology observed in the scans was annotated as *kidney* as well. A 6-fold cross-validation scheme was implemented on the 46 scans. The size of the six training sets was [38, 38, 38, 38, 39, 39] cases, resulting in test set sizes of [8, 8, 8, 8, 7, 7] cases, accordingly. Each network was trained for 20 epochs with a mini-batch size of 20 images, learning rate of 0.01 for the first 10 epochs and 0.001 for the next 10 epochs, momentum of 0.9 and weight decay of 0.0005. The preceding mini-batch size, momentum and weight decay values, were



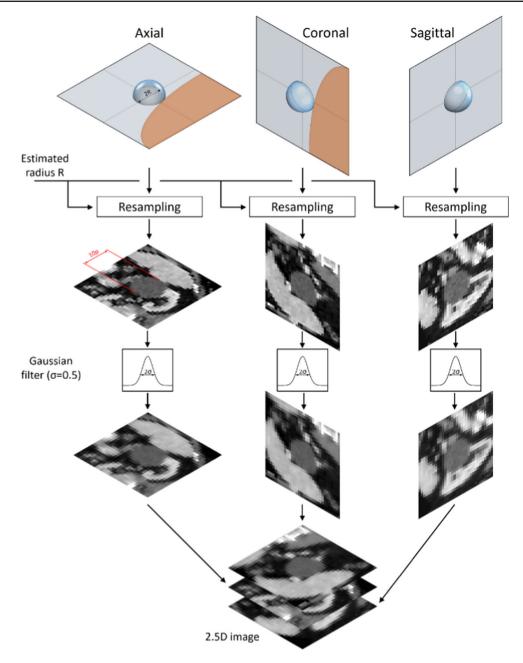


Fig. 3 The steps for constructing the 2.5D image of a cyst candidate. The candidate's 2D projections on three orthogonal planes (axial, sagittal and coronal) are used. The orthogonal slices of the CT scan are resampled according to the candidate's estimated radius so that the sus-

pected object will occupy 10×10 pixels in a 32×32 patch. The patches are then smoothed by a Gaussian filter ($\sigma = 0.5$, kernel size = 3×3) and stacked together to form a 3-channel image, also known as a 2.5D image

taken from [14]. Training was performed on an NVIDIA GTX Titan GPU, using the deep learning framework Mat-ConvNet [15]. The accuracy of the 3D segmentation was measured separately for the right and left kidneys in each CT case. Recall (blue), Precision (cyan), Intersection Over Union (IoU) [16] (yellow), and Dice coefficient [17] (red) are given for each right (top) and left (bottom) kidney in Fig. 6's bar graph. The overlap values averaged over the

dataset are further reported in Table 1. The network obtained mean dice coefficient of 0.87 for both kidneys with low std of \sim 0.095. Qualitative examples of the segmentation output and its comparison to the ground truth labeling are shown in Fig. 7. The average running time of kidney segmentation was 51 s.

Since there is no available benchmark for kidney segmentation, comparing the performance of different reported



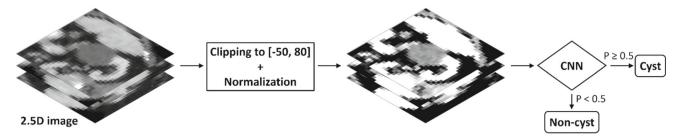
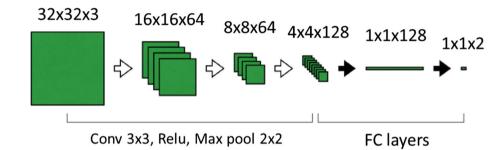


Fig. 4 The cysts candidate classification flow. The 2.5D images of the candidates are clipped to the range of [-50, 80] HU and normalized to the range of [-1, 1]. They are then fed into a CNN that computes the

probability of the candidate being a cyst. Candidates with probability larger than or equal to 0.5 are identified as cysts

Fig. 5 The cyst classification CNN architecture. The network comprises three layers of 3×3 convolution and 2×2 max pooling followed by two fully connected (FC) layers



methods is not straightforward. Different methods were evaluated on different datasets. For the most part, we have no knowledge on the quality of ground truth annotations, cyst frequency and the manner in which the presence of cysts was handled. Cuingnet et al. [18] achieved median dice coefficient of 0.96 for both kidneys using random forests for the detection of the kidneys and template deformation for segmentation. The algorithm was trained on 54 scans. Thong et al. [19] trained a CNN to determine whether the central voxel of a 2D patch belongs to the kidney. They reported median dice coefficients of 0.93 and 0.94 for the right and left kidney, respectively. Our proposed FCN is a natural extension of this idea. Several multi-organ segmentation methods reported similar dice coefficients (0.95 [20] and 0.94 [3]). These require a more comprehensive dataset, annotated for multiple internal organs.

Cyst detection

In the experiments, only cysts with a diameter larger or equal to 10 mm were considered. A different dataset was used for cyst detection, with 247 such cysts observed in 173 abdominal CT cases. 121 cases, containing 177 cysts were kept for training while the remaining 52 cases, containing 70 cysts, were used for testing. The cases were randomly selected within a one-month period. All scans were taken with contrast agent injection at the portal phase of enhancement, all have a 2.5 mm spacing between slices and pixel spacing of 0.64–0.98 mm. A line was marked along the diameter of each

cyst, allowing the extraction of its center and radius. The network was trained on a set of $\sim 12,000$ seeds extracted from 121 scans for 20 epochs with a mini-batch size of 32 images, learning rate of 0.001 for the first 10 epochs and 0.0001 for the next 10 epochs, momentum of 0.9 and weight decay of 0.001. To increase the number of cysts in the training set, we augmented the data by rotating and flipping the positive examples.

In Fig. 8, the receiver-operator curves (ROC) gives the TP rate in function of the FP rate/per patient for two major parameters: the minimum cyst seed radius (left), and the kidney adjacency (right) thresholds.

By setting the minimum cyst radius to 2.5 mm and the kidney adjacency threshold to -2.5 mm, the algorithm obtained a true-positive (TP) rate of 84.3%, meaning that 59 out of 70 cysts, larger than 10 mm, were successfully detected. Kidney segmentation was sufficiently accurate to include the required cyst candidate points in all but 2 cases. The remaining cyst miss-detections, that are 9 in number, were due to miss-classification by the CNN. The algorithm also produced 83 false detections, leading to false-positive rate of 1.6 per patient. However, 4 of these false detections were indications of severe hydronephrosis, still having clinical relevance and requiring radiologist attention.

Qualitative examples of detected cysts, missed cysts, and false-positives are shown in Fig. 9. In addition, the algorithm detected 9 more cysts smaller than 10 mm. The average running time of the cyst candidate selection and classification was 22 s.



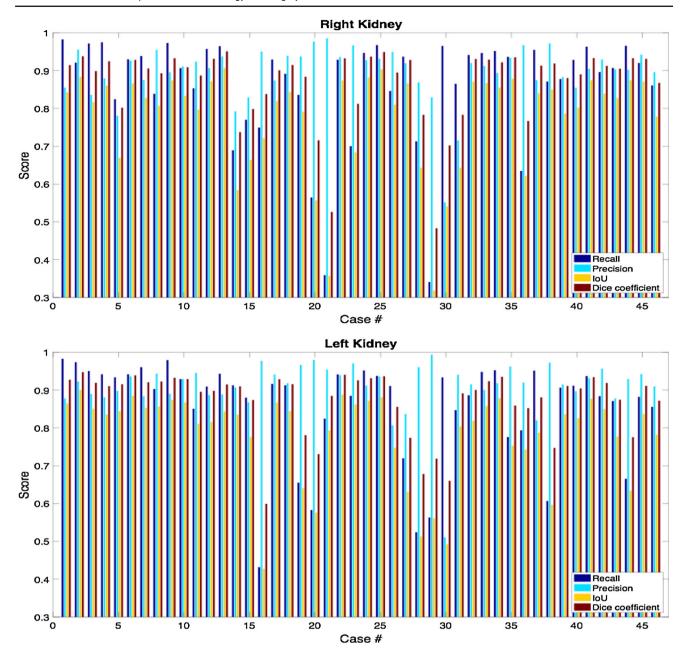


Fig. 6 Segmentation accuracy measurements for each kidney in the testing set

Discussion

We proposed a fully automatic framework for renal cysts detection, supported by a robust end-to-end segmentation of the kidneys by an FCN. The proposed kidney segmentation was successful in the presence of multiple cysts and other pathologies present in randomly selected abdominal scans, that often significantly deforms the kidney appearance. It demonstrated robust kidney segmentation and led to successful cyst detection step.

The combined 3D distance map of the kidneys and surrounding fluids provides a large number of cysts candidates

(almost 11,000 candidates for 52 scans) that were successfully handled by the classification network, outputting clinically relevant results with high sensitivity and specificity rates. In the context of automatic incidental findings detection, 1.6 FP/case means, in average, less than two unnecessary key images per case. This is an acceptable overhead given the detection benefits of the proposed system. In comparison, Liu et al. [10] reported 15 FP/case in their computer-aided system for detection of exophytic renal lesions.

In the "Identification of cyst candidates" section, we mentioned that seeds with estimated radius less than 2.5 mm are



Table 1 Average kidneys segmentation accuracy evaluated on 46 volumetric CT cases

Organ	Mean recall	Mean precision	Mean IoU	Mean dice coefficient
Right kidney	0.86 ± 0.148	0.90 ± 0.076	0.78 ± 0.137	0.87 ± 0.103
Left kidney	0.86 ± 0.136	0.91 ± 0.073	0.78 ± 0.120	0.87 ± 0.085

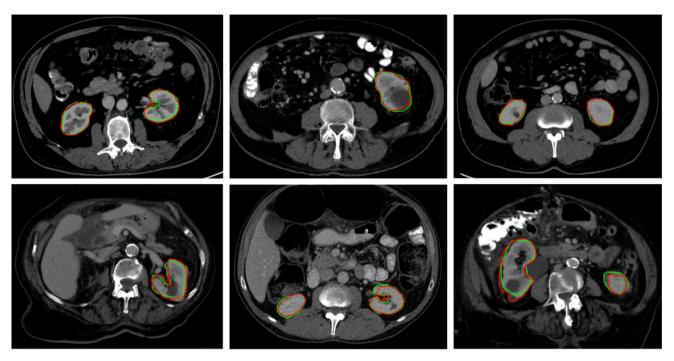


Fig. 7 Sample kidney segmentation results using the proposed FCN. The manually annotated ground truth (green) and the segmentation map produced by the FCN (red) are marked on the CT images. Simple renal

cysts are evident in all of the examples, demonstrating the network's ability to successfully segment them as part of the kidney

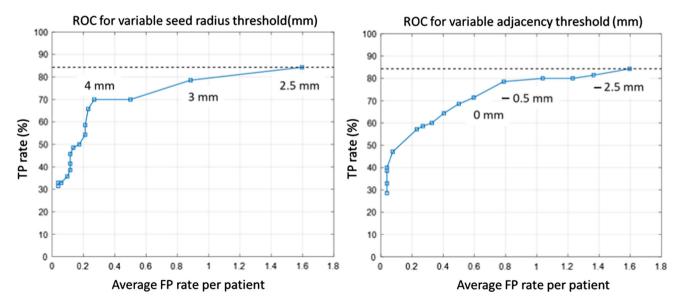


Fig. 8 Receiver-operator curves (ROC) for cyst detection: TP rate in function of the FP rate/per patient for two major parameters: the minimum cyst seed radius (left), and the kidney adjacency (right) thresholds



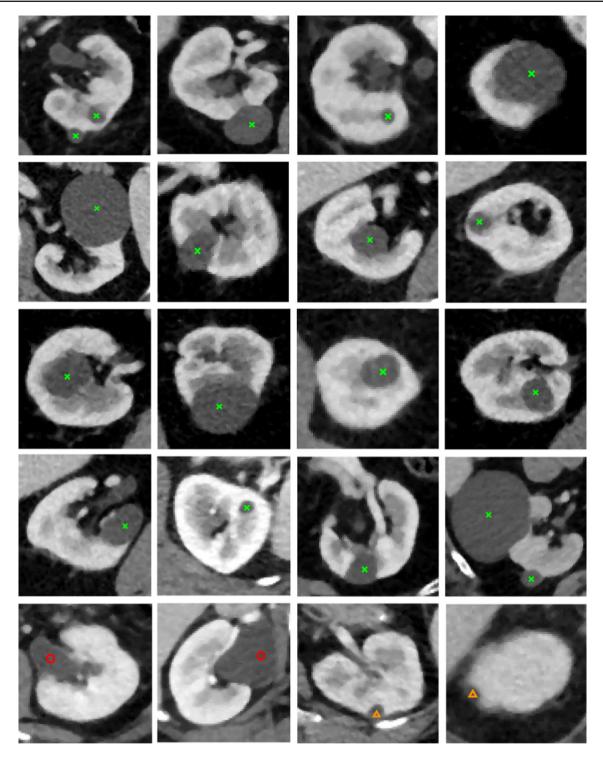


Fig. 9 Sample cyst candidates and their classification results. The candidates marked with green x are cysts that were classified correctly. The candidates marked with red circles are false-positives. The left exam-

ple is a case of hydronephrosis and the right is a case of extrarenal pelvis. The candidates marked with orange triangles are cysts that were misclassified as non-cysts



eliminated. The trade-off between the TP and FP rates can be tuned by changing the radius threshold. Setting the threshold to 3 mm significantly reduces the FP rate from average of 1.6 per patient to 0.88, but at the cost of missing seeds of 4 cysts sized 10–12 mm.

In future research the proposed method will be extended to the detection of complex cysts and, again leveraging the robust kidney segmentation algorithm, the method will be further extended to the automatic detection of renal cells carcinoma.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent For this type of study formal consent is not required.

References

- Taal MW, Chertow GM, Marsden PA, Skorecki K, Alan SL, Brenner BM (2011) Brenner and rector's the kidney. Elsevier Health Sciences, Amsterdam
- 2. Zhou X, Ito T, Takayama R, Wang S, Hara T, Fujita H (2016) Three-dimensional CT image segmentation by combining 2d fully convolutional network with 3d majority voting. In: International workshop on large-scale annotation of biomedical data and expert label synthesis. Springer, Cham, pp 111–120
- 3. Hu P, Wu F, Peng J, Bao Y, Chen F, Kong D (2016) Automatic abdominal multiorgan segmentation using deep convolutional neural network and time-implicit level sets. Int J Comput Assist Radiol Surg 12(3):399–411
- Christ PF, Elshaer MEA, Ettlinger F, Tatavarty S, Bickel M, Bilic P, Rempfler M, Armbruster M, Hofmann F, D'Anastasi M, Sommer WH (2016) Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 415–423
- Dou Q, Chen H, Jin Y, Yu L, Qin J, Heng PA (2016) 3d deeply supervised network for automatic liver segmentation from CT volumes. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp. 149–157
- Summers RM, Agcaoili CM, McAuliffe MJ, Dalal SS, Yim PJ, Choyke PL, Walther MM, Linehan WM (2001) Helical CT of von Hippel-Lindau: semi-automated segmentation of renal lesions. In: Proceedings of 2001 international conference on image processing, 2001, vol 2, pp 293–296

- Battiato S, Farinella GM, Gallo G, Garretto O, Privitera C (2009)
 Objective analysis of simple kidney cysts from CT images. In:
 IEEE International Workshop on medical measurements and applications, 2009. MeMeA 2009, pp 146–149
- Piao N, Kim JG, Park RH (2015) Segmentation of cysts in kidney and 3-D volume calculation from CT images. Int J Comput Graph Animat 5(1):1
- Badura P, Wieclawek W, Pycinski B (2016) Automatic 3D segmentation of renal cysts in CT. In: Information technologies in medicine. Springer, Cham, pp 149–163
- Liu J, Wang S, Linguraru MG, Yao J, Summers RM (2015) Computer-aided detection of exophytic renal lesions on noncontrast CT images. Med Image Anal 19(1):15–29
- Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 234–241
- Maurer CR, Qi R, Raghavan V (2003) A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. IEEE Trans Pattern Anal Mach Intell 25(2):265–270
- Roth HR, Lu L, Seff A, Cherry KM, Hoffman J, Wang S, Liu J, Turkbey E, Summers RM (2014) A new 2.5 d representation for lymph node detection using random sets of deep convolutional neural network observations. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 520–527
- Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3431

 3440
- Vedaldi A, Lenc K (2015) Matconvnet—convolutional neural networks for matlab. In: Proceeding of the ACM int. conf. on multimedia
- Jaccard P (1901) Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bulletin de la Société Vaudoise des Sciences Naturelles. 37:547–579
- 17. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26(3):297–302
- Cuingnet R, Prevost R, Lesage D, Cohen LD, Mory B, Ardon R (2012) Automatic detection and segmentation of kidneys in 3d CT images using random forests. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 66–74
- Thong W, Kadoury S, Pich'e N, Pal CJ (2016) Convolutional networks for kidney segmentation in contrast-enhanced CT scans. Comput Methods Biomech Biomed Eng Imaging Vis 1–6
- Wolz R, Chu C, Misawa K, Mori K, Rueckert D (2012) Multiorgan abdominal CT segmentation using hierarchically weighted subject-specific atlases. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 10–17

